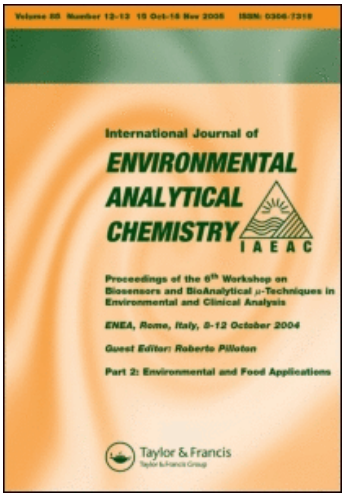


International Journal of Environmental Analytical Chemistry

Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713640455

Normalization of Sedimentary Lipid Biomarker Concentrations to Total Organic Carbon in Principal Component Analysis

Woei-Lih Jeng[a]; Bor-Cheng Han[b]
[a] Institute of Oceanography, National Taiwan University, Taipei, Taiwan [b] School of Public Health, Taipei Medical College, Taipei, Taiwan

To cite this Article Jeng, Woei-Lih and Han, Bor-Cheng(1999) 'Normalization of Sedimentary Lipid Biomarker Concentrations to Total Organic Carbon in Principal Component Analysis', International Journal of Environmental Analytical Chemistry, 73: 1, 15 — 29
To link to this Article: DOI: 10.1080/03067319908032648
URL: http://dx.doi.org/10.1080/03067319908032648

Intern. J. Environ. Anal. Chem., Vol. 73(1), pp. 15-29
Reprints available directly from the publisher
Photocopying permitted by license only

NORMALIZATION OF SEDIMENTARY LIPID BIOMARKER CONCENTRATIONS TO TOTAL ORGANIC CARBON IN PRINCIPAL COMPONENT ANALYSIS

WOEI-LIH JENG[a*] and BOR-CHENG HAN[b]

[a]*Institute of Oceanography, National Taiwan University, Taipei, Taiwan, R.O.C. and*
[b]*School of Public Health, Taipei Medical College, Taipei, Taiwan, R.O.C.*

(Received 21 January, 1998; In final form 28 August, 1998)

The objective of this study is to demonstrate the importance of normalizing lipid biomarker concentrations in sediment to total organic carbon (TOC) for principal component analysis (PCA) by using n-alkanols and aliphatic hydrocarbons in marine sediments collected from the East China Sea shelf off northern Taiwan. In performing PCA, logarithmically transformed data and z-score function transformed data along with the raw data were compared with TOC normalized data. Results show that the positions of n-alkanol variables in the loading plot using TOC normalized data are in good agreement with the organic geochemical knowledge in terms of sources. For aliphatic hydrocarbons, the positions of samples in the score plot using TOC normalized data are different from those using the raw data and z-score function transformed data. It is suggested that normalization of lipid biomarker concentrations with TOC be taken into consideration in performing PCA when the grain size distributions of sediments in a study area vary in a wide range.

Keywords: Principal component analysis; lipid biomarkers; normalization

INTRODUCTION

The multivariate method such as principal component analysis (PCA) and factor analysis is widely used for evaluating differences and observing similarities among multiple objects, especially for geochemical and environmental data.[1–6] Multivariate statistics can be a powerful means of data reduction. In performing PCA, several pre-processing procedures are generally tried. All those procedures are different mathematical operations. For geochemical samples, especially sediments, grain size plays an important role in controlling the concentrations of

* Corresponding author. Fax: +886–2–23626092. E-mail: wljeng@iodec1.oc.ntu.edu.tw.

compounds and metals.[7–10] Effects caused by grain size are almost indistinguishable from effects due to surface area; as grain size decreases, surface area increases sharply.[11] A positive linear correlation between total organic carbon (TOC) and sediment specific surface area has been demonstrated by Suess[12] and Mayer et al.[13] There are a number of ways to adjust for the grain size effect; the most common and easiest one is normalization using TOC for organic compounds since TOC, grain size and the specific surface area of sediments are interrelated. In this study, we explore (1) how the grain size effect influences the positions of the variables in the loading plots by comparing TOC normalized data with the raw data pre-processed by mathematical procedures and (2) how well the positions of variables (biomarkers) in the loading plots agree with the organic geochemical knowledge with respect to sources.

MATERIALS AND METHODS

Nine surface (top 3–4 cm) sediment samples from the East China Sea shelf off north Taiwan were collected with a box corer on board the R/V *Ocean Researcher #1* for this study (Figure. 1). Each sediment was freeze-dried and ground. Internal standards (n-$C_{24}D_{50}$ and 1-heptadecanol) were added to the sediment and extracted with benzene/methanol (1:1) in a Soxhlet apparatus for 24 h. The spiked extract was concentrated and hydrolyzed with methanolic KOH. The neutral lipids were extracted with n-hexane (4X). The nonsaponifiable lipids were subjected to silica gel (deactivated with 5% H_2O) column chromatography. Aliphatic hydrocarbons and alkanols/sterols were eluted with n-hexane and a mixture of dichloromethane/methanol (4/1, v/v), respectively. The lipids between the two fractions were removed with n-hexane/dichloromethane (2:3, v/v). The isolated alkanols/sterols were taken to dryness, redissolved in benzene, and derivatized with *N,O*-bis-(trimethylsilyl)-acetamide. Aliphatic hydrocarbons and alkanols/sterols (as TMS ethers) were analyzed by capillary gas chromatography using an HP 5890 gas chromatograph equipped with a split/splitless injector and an FID. An SGE (Australia) OCI-5 cool on-column injector was also fitted in the gas chromatograph for quantitation. Separation was achieved by an SE-30 capillary column (30 m × 0.25 mm i.d.). Oven temperature programming was 45–90°C at 15°C/min and 90–280°C at 3°C/min for analyzing aliphatic hydrocarbons, and 45–90°C at 15°C/min, 90–270°C at 3°C/min, 20 min at 270°C, 270–280°C at 10°C/min, and 20 min at 280°C for analyzing alkanols/sterols. GC traces for aliphatic hydrocarbons and alkanols/sterols are the same as those given elsewhere.[14] The relative precision of the lipid determination was estimated to be 2~8%.

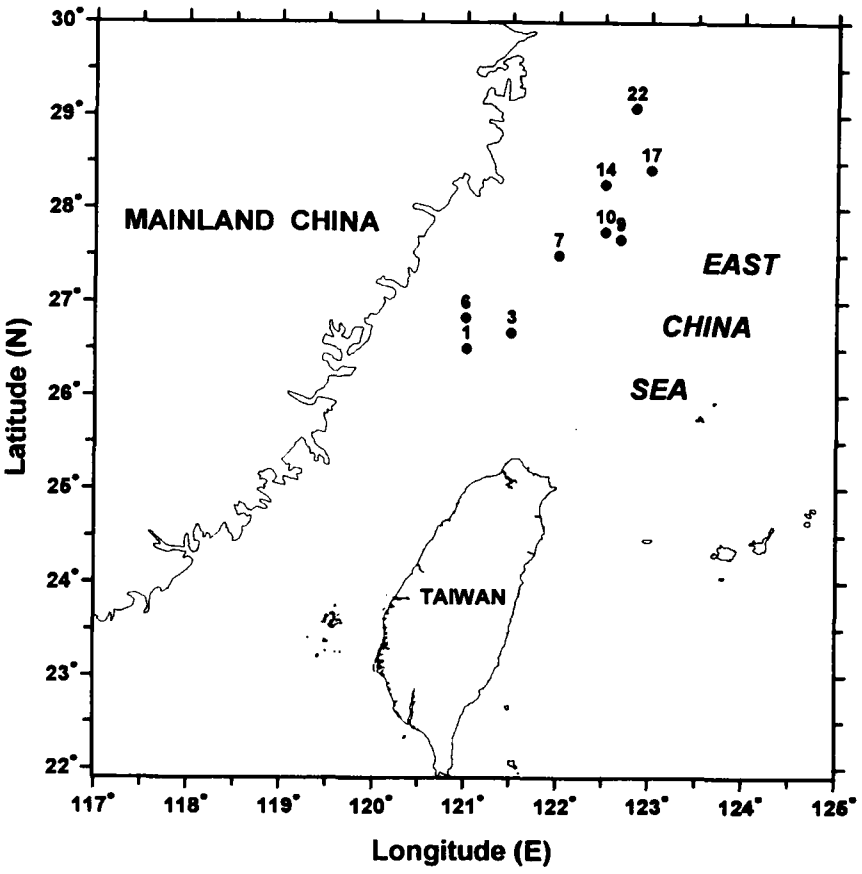


FIGURE 1 Location of sampling sites on the East China Sea shelf off north Taiwan

Sediment samples were oven-dried in air for 48 h. Total organic carbon was determined by the dichromate-acid oxidation method[15] modified by addition of Ag_2SO_4 to H_2SO_4 at the rate of 15g/L. Titration was carried out with an automatic titrator (Metrohm 702 SM Titrino, Switzerland). The relative standard deviation of TOC determination was generally < 1%.

Data pre-processing includes the use of logarithmical transformation and z-score function transformation in which data were centered by subtracting the means of the variables and scaled by dividing by the standard deviations of the variables. The raw data and TOC normalized data were included for comparison. PCA was performed with the xlSTAT data analysis toolbox, version 2.6.[16]

RESULTS AND DISCUSSION

In PCA, the plots of loadings indicate relationships among the variables, and the plots of scores give the positions of the samples in the co-ordinates of the principal components.[17] Relationships among variables can be visually inspected from loading plots. In the present study, variables are the lipid biomarkers. n-Alkanols of biological origins are dominantly even-carbon-numbered; therefore, n-C_{14}, n-C_{16}, n-C_{18}, n-C_{20}, n-C_{22}, n-C_{24}, n-C_{26} and n-C_{28} alkanols are chosen as variables. Concentrations of even-carbon-numbered n-alkanols and TOC are listed in Table I. Plots of loadings using the raw data, logarithmically transformed data, z-score function transformed data, and TOC normalized data are given in Figure 2a-d, respectively. In Figure 2a-c, the first two principal components loading plots indicate two groups of variables: n-C_{20}, n-C_{22}, n-C_{24}, n-C_{26}, and n-C_{28} alkanols, which are typical of higher plant waxes[18,19] and hence are considered from terrigenous sources, and n-C_{14}, n-C_{16} and n-C_{18} alkanols, which may be formed by hydrolysis of esterified alcohols derived from a wide variety of organisms such as zooplankton[20] and are considered from non-terrigenous sources although these eight n-alkanols behave similarly. It is also noted that the variables of terrigenous sources are correlated and those of non-terrigenous sources form a cluster. However, it is seen from Figure 2d that the variables of non-terrigenous sources form a group and those of terrigenous sources form another. The correlation between variables can be estimated from the vectors which are the connecting lines from the origin to the variables (not shown). Vectors which point in the same direction indicate a high positive correlation between variables, whereas negative correlations are indicated by vectors pointing in opposite directions, and orthogonal vectors (those meeting at 90°) indicating no correlation. In Figure 2d the resultant vector of non-terrigenous sources is roughly along PC-2 if plotted, and that of terrigenous sources is roughly along PC-1 if plotted. PC-1 and PC-2 are mutually perpendicular as is defined in PCA. This means that the vectors of non-terrigenous sources are roughly orthogonal to those of terrigenous sources, indicating no correlation between non-terrigenous and terrigenous sources, i.e., independent of each other, that is in good agreement with the organic geochemical knowledge. The difference between the loading plot using TOC normalized data and those using the other three data sets is apparently due to the grain size effect because TOC ranges from 0.18 to 0.66 g/100 g (Table I). Therefore, using TOC normalized data for plots of scores may be more accurate and precise for source interpretation of samples.

TABLE I Concentrations of even carbon n-alkanols (ng/g) and total organic carbon (g/100 g) in marine sediments off north Taiwan

Sample	$C_{14}OH$	$C_{16}OH$	$C_{18}OH$	$C_{20}OH$	$C_{22}OH$	$C_{24}OH$	$C_{26}OH$	$C_{28}OH$	*TOC*
1	29	71	59	64	230	159	170	149	0.48
3	12	81	61	60	234	144	107	119	0.53
6	31	69	77	77	285	273	304	306	0.66
7	33	130	75	85	320	235	322	323	0.65
9	13	42	22	25	88	53	44	55	0.18
10	24	70	46	23	47	18	9	12	0.24
14	112	113	92	77	241	199	244	211	0.46
17	22	73	66	74	272	177	221	238	0.5 1
22	68	162	104	115	288	261	378	302	0.65

TABLE II Concentrations of aliphatic hydrocarbons (ng/g) and total organic carbon (g/100 g) in marine sediments off north Taiwan

Sample	*pristane*	C_{17}	C_{19}	C_{21}	C_{25}	C_{27}	C_{29}	C_{31}	*TOC*
1	3	6	35	70	95	145	370	345	0.48
3	10	11	37	73	102	153	353	306	0.53
6	5	0	16	43	97	146	332	428	0.66
7	7	4	30	65	104	176	449	469	0.65
9	0	0	8	15	24	38	96	90	0.18
10	11	18	16	21	31	42	102	121	0.24
14	6	4	21	49	96	158	374	447	0.46
17	0	0	22	50	77	119	289	331	0.51
22	89	52	55	104	105	190	538	559	0.65

TABLE III Concentrations of sterols (ng/g) and total organic carbon (g/100 g) in marine sediments off north Taiwan

Sample	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*TOC*
1	216	409	332	161	187	233	369	0.48
3	257	507	419	179	228	287	458	0.53
6	196	353	307	141	163	226	353	0.66
7	487	1090	847	374	424	554	982	0.65
9	128	226	185	86	90	128	202	0.18
10	100	326	192	70	83	147	207	0.24
14	337	927	579	233	284	307	559	0.46
17	449	1200	697	281	346	434	776	0.51
22	776	2230	1770	822	905	736	1600	0.65

Sterol identifications: 1, 22-dehydrocholesterol; 2, cholesterol; 3, diatomsterol; 4, 24-methylenecholesterol ; 5, campesterol; 6, stigmasterol; 7, β-sitosterol.

The major n-alkanes found in marine phytoplankton[21] and benthic algae[22] are n-C_{15} and n-C_{17}. Pristane is a well known biomarker of zooplankton, particularly of calanoid copepods.[23] Therefore, n-C_{17} and pristane are chosen for marine sources. Major plant wax n-C_{25}, n-C_{27}, n-C_{29} and n-C_{31} alkanes are used to represent terrigenous sources.[24] Concentrations of pristane and odd-carbon-numbered n-alkanes are listed in Table II. The loading plots of these alkanes for four data sets show that the variables have very similar positions (Figure 3a-d) and that the vectors of the variables of marine sources (n-C_{17} and pristane) are roughly orthogonal to those of terrigenous sources (n-C_{25}, n-C_{27}, n-C_{29} and n-C_{31}) if plotted, indicating that the two sources have no correlation and hence in agreement with the organic geochemical knowledge. However, a close examination indicates that the positions of variables in the loading plot using TOC normalized data exhibit a bit wider spread than those using the other three data sets. This shows little grain size effect on pristane and n-alkanes in the sediments. Furthermore, the positions of samples in the score plot using TOC normalized data are quite different from those using the raw data and z-score function transformed data. For instance, the former shows scatter sample points (Figure 4d), whereas the latter form three groups of samples (Figure 4a and c).

Using the sterol data given in Table III for PCA, all the variables of marine sources (22-dehydrocholesterol, cholesterol, diatomsterol, and 24-methylenecholesterol[25]) and terrigenous sources (campesterol, stigmasterol, and β-sitosterol[26]) in the plots of loadings are closely clustered and correlated because PC-1 accounts for 96–98% of the variance (Figure 5a-d). As PC-2 has much lower eigenvalues than PC-1, the separation is not distinct. Another possible cause is that the sterols representing terrigenous inputs such as campesterol, stigmasterol and β-sitosterol could also originate from algae.[27,28] It is noted that from Figure 5 the positions of variables in the loading plots using raw data and TOC normalized data exhibit a little wider spread than those using the other two data sets. Moreover, a slight difference can be seen from the positions of variables. Figure 5d shows that 22-dehydrocholesterol (#1) and cholesterol (#2) generally considered from zooplankton[25,29] form a group and that diatomsterol (#3) and 24-methylenecholesterol (#4) generally considered from diatoms[25,30] form another. However, this is not seen in Figure 5a.

In summary, the present results indicate that the grain size effect appears to play a role in the plot of loading. It is suggested that, in pre-processing organic geochemical data for PCA, normalization of lipid biomarker concentrations with TOC be taken into consideration when the grain size distributions of sediments in a study area vary in a wide range.

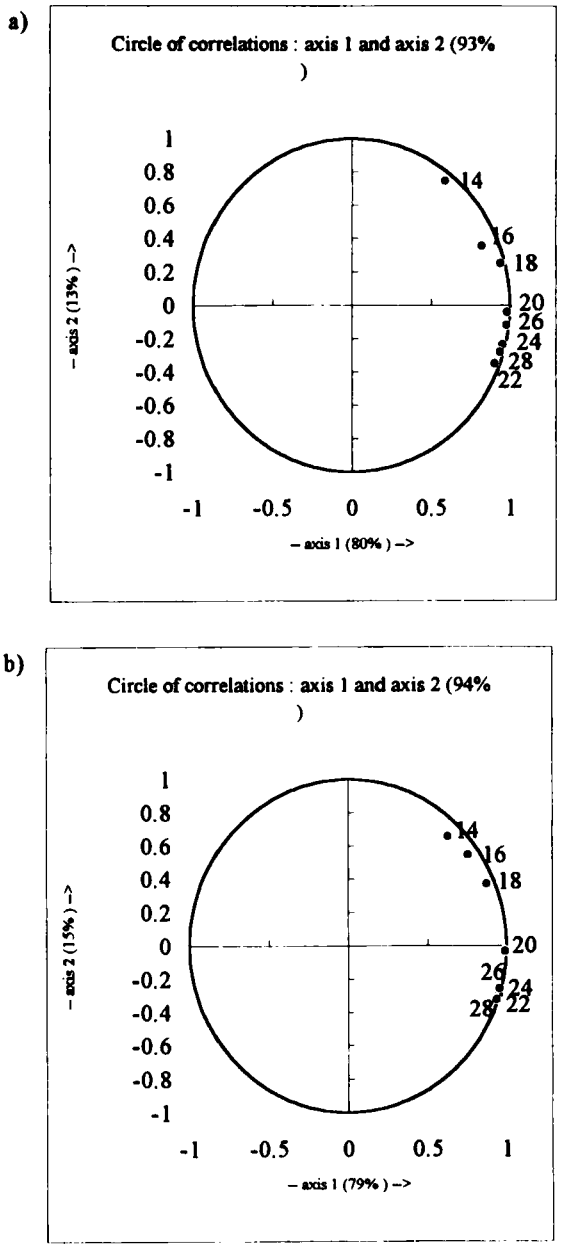
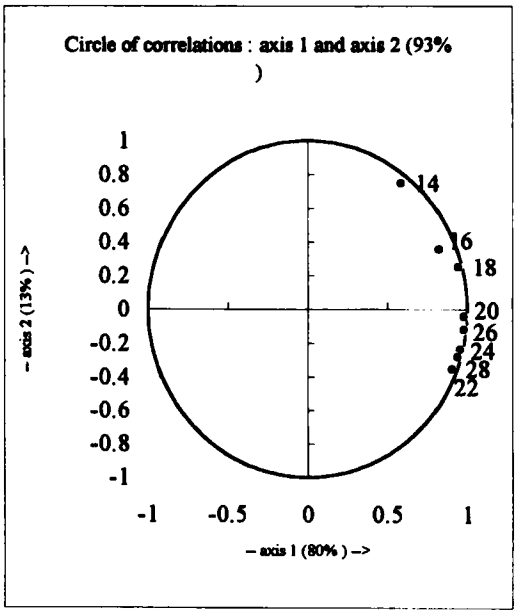
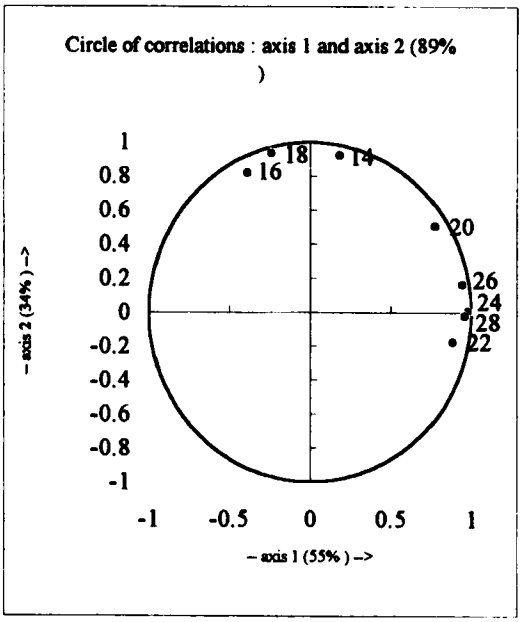


FIGURE 2 n-Alkanols in marine sediments, plots of loadings on the first two principal components (PC- 1 and PC-2) by using (a) the raw data, (b) logarithmically transformed data, (c) z-score function transformed data, and (d) TOC normalized data. Numbers indicate the carbon number of n-alkanols

c)



d)



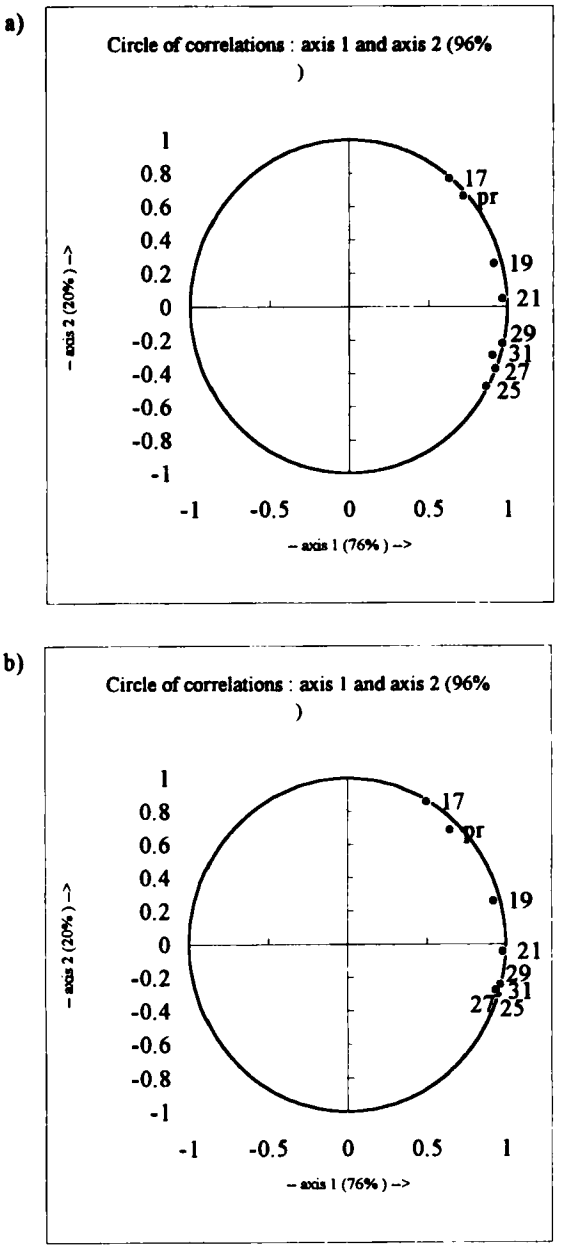
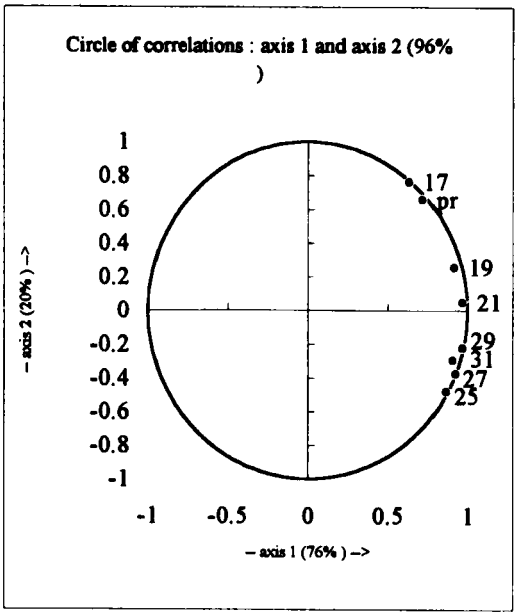
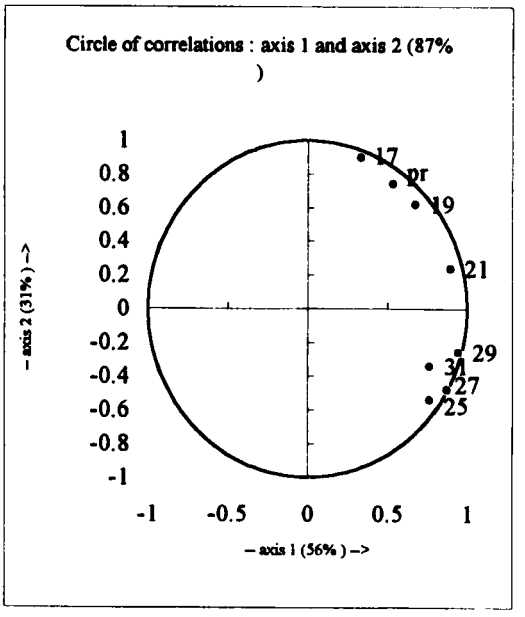


FIGURE 3 Aliphatic hydrocarbons in marine sediments, plots of loadings on the first two principal components (PC-1 and PC-2) by using (a) the raw data, (b) logarithmically transformed data, (c) z-score function transformed data, and (d) TOC normalized data. Numbers indicate the carbon number of n-alkanes, and pr denotes pristane

c)



d)



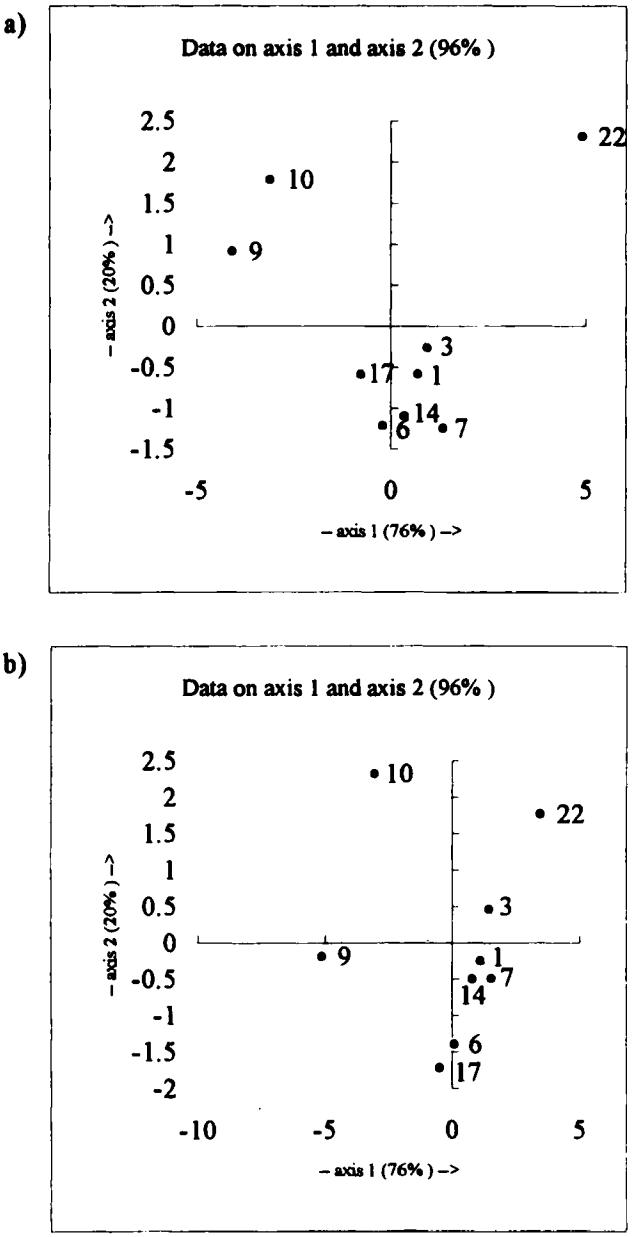
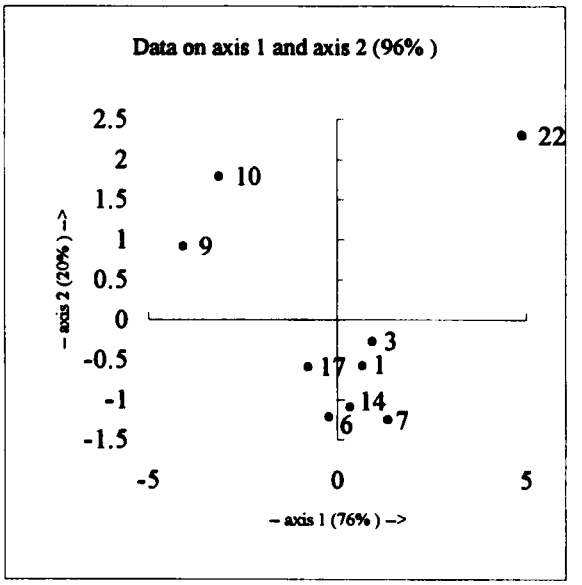
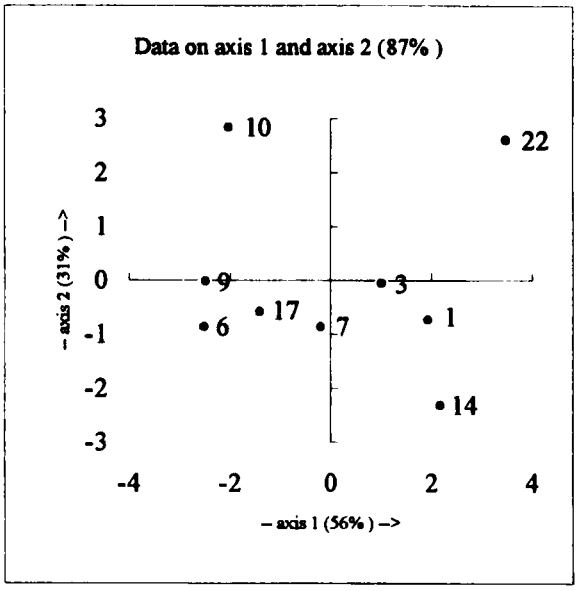


FIGURE 4 Aliphatic hydrocarbons in marine sediments, plots of scores in the co-ordinates of first two principal components (PC-1 and PC-2) by using (a) the raw data, (b) logarithmically transformed data, (c) z-score function transformed data, and (d) TOC normalized data. Numbers correspond to sampling sites shown in Fig. 1

c)



d)



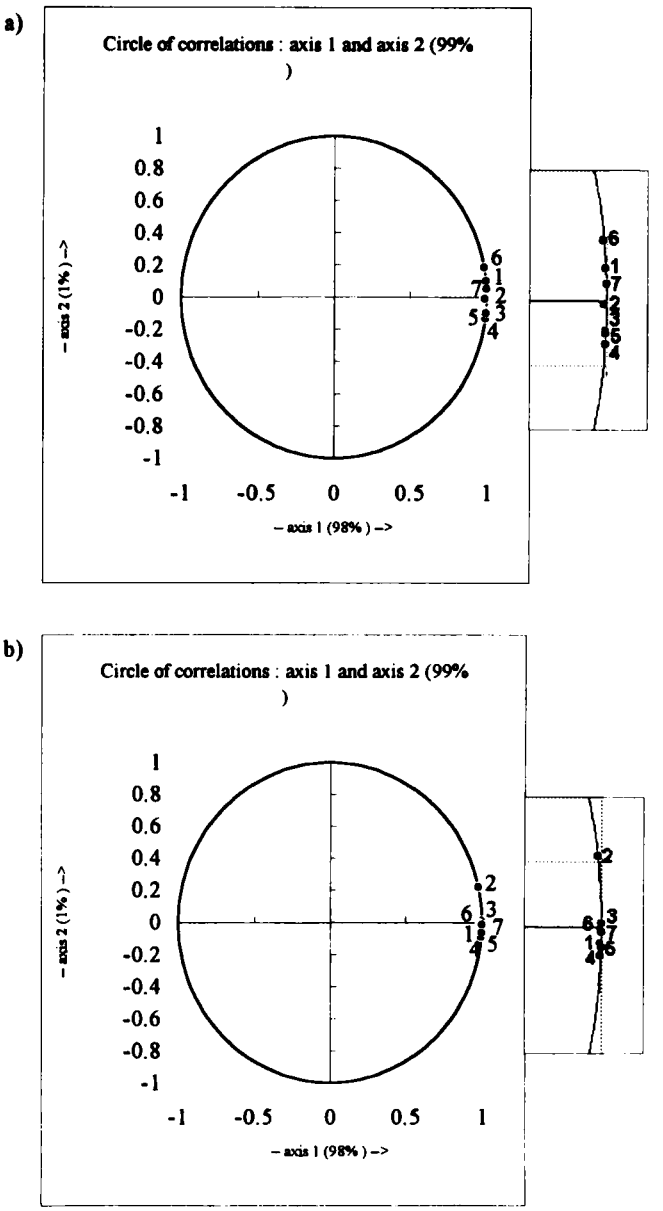
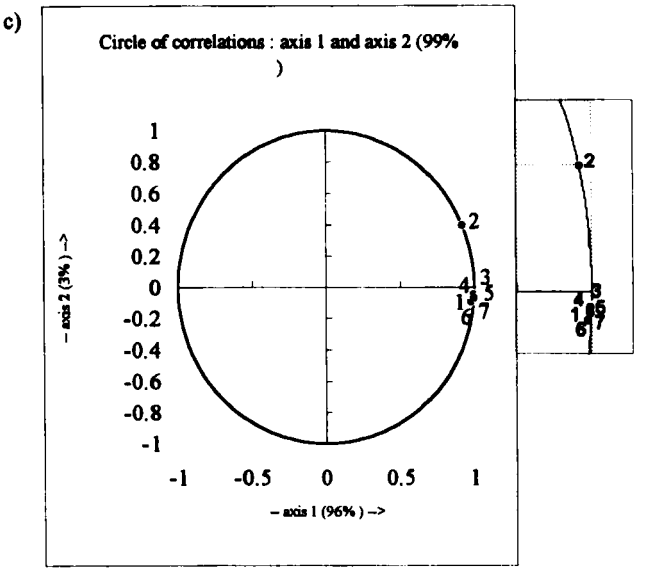
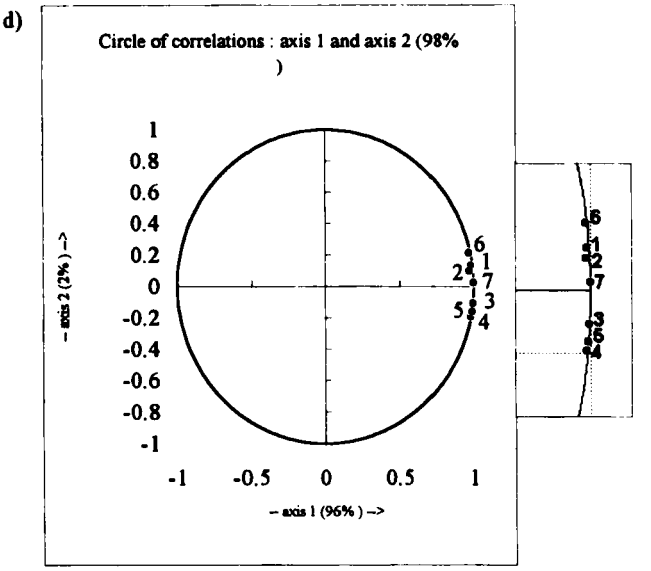


FIGURE 5 Sterols in marine sediments, plots of loadings on the first two principal components (PC-1 and PC-2) by using (a) the raw data, (b) logarithmically transformed data, (c) z-score function transformed data, and (d) TOC normalized data. Numbers indicate sterols: 1, 22-dehydrocholesterol; 2, cholesterol; 3, diatomsterol; 4, 24-methylenecholesterol ; 5, campesterol; 6, stigmasterol, and 7, β-sitosterol. Expanded regions for crowded variables are shown





Acknowledgements

We would like to thank the captain, crew and technicians of R/V *Ocean Researcher #1* for help with sample collection. Our thanks also go to one anonymous reviewer for suggestions. This study was financially supported by the National Science Council of the R.O.C. (grant # NSC86–2611-M002A-011-K2).

References

[1] A. Grant, *Mar. Pollut. Bull.* **21**. 297–299 (1990).

[2] F. I. Onuska, and S. Davies, *Intern. J. Environ. Anal. Chem.* **43**, 137–150 (1991).

[3] J. O. Grimalt, and J. Olive, *Anal. Chim. Acta* **278**, 159–176 (1993).

[4] T. A. T. Aboul-Kassim, and B. R. T. Simoneit, *Mar. Pollut. Bull.* **30**, 63–73 (1995).

[5] S. M. Mudge, and C. E. Norris, *Mar. Chem.* **57**, 61–84 (1997).

[6] J. C. Colombo, N. Silverberg, and J. N. Gearing, *Org. Geochem.* **25**, 211–225 (1997).

[7] C. Gonzalez, A. Saliot, and P. Pillon, *Intern. J. Environ. Anal. Chem.* **22**, 47–59 (1985).

[8] A. Barouxis, P. Scribe, J. Dagaut, and A. Saliot, *Org. Geochem.* **13**, 773–783 (1988).

[9] D. H. Loring, *Mar. Chem.* **29**, 155–168 (1990).

[10] A. Grant and R. Middleton, *Estuar. Coast. Shelf Sci.* **31**, 71–85 (1990).

[11] A. J. Horowitz, *A Primer on Trace Metal-Sediment Chemistry*. U.S. Geological Survey Water-Supply Paper 2277 (1985).

[12] E. Suess, *Geochim. Cosmochim. Acta* **37**, 2435–2447 (1973).

[13] L. M. Mayer, P. T. Rahaim, W. Guerin, S. A. Macko, L. Watling, and F. E. Anderson, *Estuar. Coast. Shelf Sci.* **20**, 491–503 (1985).

[14] W. L. Jeng, and M. P. Chen, *Org. Geochem.* **23**, 301–310 (1995).

[15] H. E. Gaudette, W. R. Flight, L. Toner, and D. W. Folger, *J. Sedi. Petrol.* **44**, 249–253 (1974).

[16] T. Fahmy, xlSTAT (version 2.6), Paris, France (1997).

[17] V. Zitko, *Mar. Pollut. Bull.* **28**, 718–722 (1994).

[18] A. P. Tulloch, in: *Chemistry and Biochemistry of Natural Waxes* (P. E. Kollatukudy, ed. Elsevier, New York, 1976) pp. 235–287.

[19] J. Vioque, J. Pastor, and E. Vioque, *J. Am. Oil. Chem. Soc.* **71**, 671–673 (1994).

[20] J. J. Boon and J. W. de Leeuw, *Mar. Chem.* **7**, 117–132 (1979).

[21] M. Blumer, R. R. L. Guillard, and T. Chase, *Mar. Biol.* **8**, 183–189 (1971).

[22] W. W. Youngblood, M. Blumer, R. L. Guillard, and F. Fiore, *Mar. Biol.* **8**, 190–201 (1971).

[23] M. Blumer, M. M. Mullin, and D. W. Thomas, *Science* **140**, 974 (1963).

[24] G. Eglinton and R. J. Hamilton, in: *Chemical Plant Taxonomy* (T. Swain, ed. Academic Press, New York, 1963) pp. 187–196.

[25] R. B. Gagosian, J. K. Volkman, and G. E. Nigrelli, in: *Advances in Organic Geochemistry 1981* (M. Bjoroy et al., eds. Wiley, Chichester, 1983) pp. 369–379.

[26] L. J. Goad, in: *Lipids and Lipid Polymers in Higher Plants* (M. Tevini and H. K. Lichtenthaler, eds. Springer Verlag, 1977) pp. 146–168.

[27] J. K. Volkman, *Org. Geochem.* **9**, 83–99 (1986).

[28] J. de Leeuw and M. Baass, in: *Biological Markers in the Sedimentary Record* (R. B. Johns, ed. Elsevier, Amsterdam, 1986) pp. 101–123.

[29] S. G. Wakeham, *Geochim. Cosmochim. Acta* **51**, 3051–3069 (1987).

[30] H. G. Harvey, *Deep-Sea Res. II* **41**, 783–796 (1994).